



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 164274

TO: Richard Schnizer
Location: rem/2D30/2C18
Art Unit: 1635
Thursday, September 08, 2005
Case Serial Number: 09/910432

From: Barb O'Bryen
Location: Biotech-Chem Library
Remsen 1a69
Phone: 571-272-2518

BOB
barbara.obryen@uspto.gov

Search Notes

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

^o
This Page Blank (uspto)

RA EU Arabidopsis sequencing project;
 RN Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 [3]

RA SEQUENCE FROM N.A.
 RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
 RA Maves H.W., Lemcke K., Mayer K.F.X.;
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [4]

RA SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL022141; CAAL18129.1; -

DR EMBL; AL161589; CAB80294.1; -

DR PIR; T04592; T04592.

DR InterPro; IPR002952; Eggshell.

DR PRINTS; PR01228; EGGSHELL.

SQ SEQUENCE 221 AA; 20344 MW; D88DAE2B2C57A9F8 CRC64;

Query Match 74.8%; Score 223; DB 2; Length 221;

Best Local Similarity 55.8%; Pred. No. 6.1e-11;

Matches 43; Conservative 1; Mismatches 7; Indels 26; Gaps 1;

Qy 1 GGGGGGGG-----GGGGGGGGGGRRDRRRRRGGG 34

Db 96 GGGGGGGGGGSSGGGGGGGNGKDNKSHRNKSSGGGGGGGGGNGSGNGSGRGGG 155

Qy 35 GGGGGGGGGGGGGGGG 51

Db 156 GGGGGGGGGGGGGGGG 172

RESULT 3

ID Q85215 PRELIMINARY; PRT; 329 AA.

AC Q85215;

DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Hypothetical protein OSJNB0078A17.2.

GN Name:OSJNB0078A17.2;

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,

RA Overton II L.L., Taitrin T., Kim M.M., Bera J.J., Jin S.S.,

RA Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Haiao J., Blunt S.,

RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,

RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,

RA White O., Salzberg S.L., Frazer C.M.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Buell R.;

RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC091532; AA038820.1; -

DR Gramene; Q85215; -

KW Hypothetical protein.

SQ SEQUENCE 329 AA; 33756 MW; 0530227C1F442C71 CRC64;

Query Match 70.3%; Score 209.5; DB 2; Length 329;

Best Local Similarity 75.5%; Pred. No. 1e-09;

Matches 40; Conservative 0; Mismatches 4; Indels 9; Gaps 2;

Qy 1 GGGGGGGGGGGG--GGGGGGGGRRDRRRRRGGGGGGGGGGGGGGG 51

Db 148 GGGGGGGGGGGDVGGGGGGGDS-----NVGGGGGGGGGGGGGGG 193

RESULT 4

RT "The genome sequence of Drosophila melanogaster.";

CAZ_DROME
 ID -CAZ_DROME STANDARD; PRT; 399 AA.
 AC Q27294; Q24445; Q8T3M2; Q9VXI4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE RNA-binding protein cageza (Sarcoma-associated RNA-binding fly homolog) (P19).
 DE NCBI_TaxID=7227;
 GN Name=cag; Synonym=SAARFH; ORFName=CG3606;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS A AND B), FUNCTION, SUBCELLULAR LOCATION, DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
 RP STRAIN=Canton-S;
 RX MEDLINE=95349623; PubMed=7623847;
 RA Immanuel D., Zinszner H., Ron D.;
 RT "Association of SARFH (sarcoma-associated RNA-binding fly homolog) with regions of chromatin transcribed by RNA polymerase II.";
 RL Mol. Cell. Biol. 15:4562-4571(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RP STRAIN=Canton-S;
 RX MEDLINE=95223793; PubMed=7708500;
 RA Stelow D.T., Haynes S.R.;
 RT "Cageza, a Drosophila gene encoding a novel RNA binding protein, shares homology with EWS and TUS, two genes involved in human sarcoma formation.";
 RL Nucleic Acids Res. 23:835-843(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos J.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S.-Y., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";

Science 287:2185-2195(2000).
 [4]
 GENOME REANNOTATION, AND ALTERNATIVE SPLICING.
 MEDLINE=22426069; PubMed=12537572;
 Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 Lewis S.E.;
 "Annotation of the Drosophila melanogaster euchromatic genome: a
 systematic review";
 Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [5]
 SEQUENCE FROM N.A. (ISOFORM B).
 STRAIN=Berkeley; TISSUE=Embryo;
 Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.W.,
 Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 George R.A., Gonzalez M., Guarin H., Kronmiller B., Li P.W., Liao G.,
 Miranda A., Mungall C.J., Nunoo J., Pacleb J.M., Faragas V., Park S.,
 Patel S., Phouanavong S., Wan K.H., Yu C., Lewis S.E., Rubin G.M.,
 Celniker S.E.;
 Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
 [6]
 SEQUENCE OF 39-404 FROM N.A.
 STRAIN=Oregon-R;
 Haynes S.R.;
 Submitted (APR-1988) to the EMBL/GenBank/DBJ databases.
 [7]
 SEQUENCE OF 212-261 FROM N.A.
 STRAIN=Oregon-R;
 MEDLINE=8717568; PubMed=3031652;
 Haynes S.R., Rebert M.L., Moxer B.A., Forquignon F., Dawid I.B.;
 "Pen repeat sequences are GGN clusters and encode a glycine-rich
 domain in a Drosophila cDNA homologous to the rat helix destabilizing
 protein";
 Proc. Natl. Acad. Sci. U.S.A. 84:1819-1823(1987).
 CC -!- FUNCTION: May participate in a function common to the expression
 of most genes transcribed by RNA polymerase II.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Comment=Additional isoforms seem to exist;
 CC Name=B;
 CC IsoId=Q27294-1; Sequence=Displayed;
 CC Name=A;
 CC IsoId=Q27294-2; Sequence=VSP_005778;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Ubiquitous. Enriched in the brain and central
 nervous system during embryogenesis. Enriched in the adult head.
 CC Embryos contain both isoforms A and B, whereas later in
 development (heads and torsos) only isoform B is detected.
 CC -!- DEVELOPMENTAL STAGE: Expressed in the developing embryo from the
 earliest stages of cellularization and is subsequently found in
 many cell types.
 CC -!- MISCELLANEOUS: 'Cabeza' means 'head' in Spanish.
 CC -!- SIMILARITY: Belongs to the RNP TET family.
 CC -!- SIMILARITY: Contains 1 RanBP2-type zinc finger.
 CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U13178; AAA86955.1; -;
 CC EMBL; L37083; AAC41563.1; -;
 CC EMBL; AE003501; AAF48578.2; -;
 CC EMBL; AE003501; AAN09389.1; -;

EMBL; AY094763; AAM11116.1; ALT_SEQ.
 DR EMBL; BT004875; AAC45231.1; -;
 DR EMBL; M15765; AAA70425.1; -;
 DR PIR; S54729; S54729.
 DR HSP; O95218; INOZ.
 DR FlyBase; FBgn0011571; caz.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR InterPro; IPR001876; Znf_RanGDP.
 DR Pfam; PF00076; RRM_1; 1.
 DR Pfam; PF00641; zf-RanBP; 1.
 DR SMART; SM00360; RRM; 1.
 DR SMART; SM00547; Znf RBZ; 1.
 DR PROSITE; PS01012; RRM; 1.
 DR PROSITE; PS01358; ZF_RANBP2_1; 1.
 DR PROSITE; PS01099; ZF_RANBP2_2; 1.
 KW Alternative splicing; Metal-binding; Nuclear protein; RNA-binding;
 KW Zinc-finger.
 FT DOMAIN 4 393 Gly-rich.
 FT DOMAIN 120 206 RNA-binding (RRM).
 FT ZN_FING 275 304 RanBP2-type.
 FT VARSPDIC 4 47 Missing (in isoform A).
 FT CONFLICT 39 41 PNY -> LFI (in Ref. 4).
 FT CONFLICT 92 92 H -> P (in Ref. 1, 2 and 6).
 FT CONFLICT 109 109 Missing (in Ref. 1, 2 and 6).
 FT CONFLICT 254 254 G -> GNGGGG (in Ref. 1, 2 and 6).
 FT CONFLICT 278 278 D -> E (in Ref. 6 and 7).
 FT CONFLICT 384 393 DGGPMRDGG -> MVDQKRWS (in Ref. 6).
 SQ SEQUENCE 399 AA; 38839 MW; 3001FC96BDD6FDCA CRC64;
 Query Match 69.3%; Score 206.5; DB 1; Length 399;
 Best Local Similarity 68.9%; Pred. No. 2:1e-09;
 Matches 42; Conservative 0; Mismatches 2; Indels 17; Gaps 3;
 Qy 1 GGGGGGGG-----GGGGGGGGGGGGRDRRRRRRRGGGGGGG---GGGGGGGGGG 48
 Db 215 GGGGGGGGGGFGGRRGGGGGGGGGGGGGGRFD-----RGGGGGGRYDRGGGGGGGGG 269
 Qy 49 G 49
 Db 270 G 270
 RESULT 5
 ID Q18444 PRELIMINARY; PRT; 136 AA.
 AC Q18444;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein C34D4.11.
 GN Name=C34D4.11; ORFNames=C34D4.11;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxId=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=9009613; PubMed=9851916;
 RG WormBase Consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Du Z., Le T.T.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;


```

RA Zhou L., Fraser C.W.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000010; AAU48078.1; -.
SQ SEQUENCE 199 AA; 19605 MW; 39022C03692C8D89 CRC64;

Query Match 68.1%; Score 203; DB 2; Length 199;
Best Local Similarity 68.9%; Pred. No. 2.2e-09;
Matches 42; Conservative 0; Mismatches 9; Indels 10; Gaps 3;

Qy 1 GGGGGGGGGGGG--GGGGGGGRRDRRRRRRGGGGG--GGGG-----GGGGGGGGGG 50
Db 116 GGGGGGGGGGGGDDGGVGGGGGGVGGGRDWRGGGGGGGAGAGRSGGGGGGGGGG 175

Qy 51 G 51
Db 176 G 176

RESULT 9
Q9U2U0 PRELIMINARY; PRT; 285 AA.
AC Q9U2U0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Y116A8C.35.
GN Name=uaf-2; ORFNames=Y116A8C.35;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=95069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurray A.A.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL117204; CAB55137.1; -.
DR PIR; T31503; T31503.
DR HSP; Q01081; 1UNT.
DR InAct; Q9U2U0; -.
DR WormBase; WBGene0000698; uaf-2.
DR WormPep; Y116A8C.35; C233341.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR009106; C:ART_C.
DR InterPro; IPR000504; RNA_rec_mot.
DR InterPro; IPR009145; U2_small.
DR InterPro; IPR000571; Znf_CCCH.
DR Pfam; PF00076; RRM 1; 1.
DR Pfam; PF00642; Zf-CCCH; 2.
DR PRINTS; PR01848; UZAUFACTOR.
DR SMART; SM00360; RRM; 1.
DR SMART; SM00356; Znf_C3H1; 2.
DR PROSITE; PS0102; RRM; 1.
KW Hypothetical protein.
SQ SEQUENCE 285 AA; 30856 MW; E2B932629653C040 CRC64;

Query Match 68.1%; Score 203; DB 2; Length 285;
Best Local Similarity 76.5%; Pred. No. 3e-09;
Matches 39; Conservative 0; Mismatches 8; Indels 4; Gaps 1;

Qy 1 GGGGGGGGGGGG--GGGGGGGRRDRRRRRRGGGGGGGGGGGGGGGGGGGGGGGGGG 51
Db 218 GGGGGGGGGGGG--GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 264

```

RESULT 10

```

Q8QKX8 PRELIMINARY; PRT; 698 AA.
AC Q8QKX8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE EsV-1-144.
GN Name=ORF 144;
OS Ectocarpus siliculosus virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX NCBI_TaxID=37665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EsV-1;
RX MEDLINE=20192171; PubMed=10725207; DOI=10.1006/viro.2000.0225;
RA Delaroque N., Wolf S., Muller D.G., Knippers R.;
RT "Characterization and immunolocalization of major structural proteins
RT in the brown algal virus EsV-1.";
RL Virology 269:148-155(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=EsV-1;
RA Delaroque N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF204951; AAP28325.1; -.
SQ SEQUENCE 698 AA; 74035 MW; C7638DDB922E0D50 CRC64;

Query Match 67.8%; Score 201.5; DB 2; Length 698;
Best Local Similarity 68.4%; Pred. No. 8.2e-09;
Matches 39; Conservative 0; Mismatches 3; Indels 15; Gaps 2;

Qy 1 GGGGGGGGGG-----GGGGGGGGGRRDRRRRRRGGGGGGGGGGGGGGGGGGGG 51
Db 454 GGGGGGGGTGGAGSGAGSGGGGGGTGGAG-----GGGGGGGGGGGGGGGGGG 501

```

RESULT 11

```

Q95UW6 PRELIMINARY; PRT; 161 AA.
AC Q95UW6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE No on or off transient A (Fragment).
GN Name=nonA;
OS Drosophila ezoana (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=47313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E20;
RA Huttunen S., Vieira J., Hoikkala A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY012600; AAG48870.1; -.
DR FlyBase; FBgn0046740; Dezo\nonA.
FT NON TER 1 1
FT NON TER 161 161
SQ SEQUENCE 161 AA; 15708 MW; 28A68346D17AB988 CRC64;

Query Match 67.4%; Score 201; DB 2; Length 161;
Best Local Similarity 77.1%; Pred. No. 2.7e-09;
Matches 37; Conservative 0; Mismatches 7; Indels 4; Gaps 1;

Qy 1 GGGGGGGGGGGG--GGGGGGGRRDRRRRRRGGGGGGGGGGGGGGGGGGGGGGGGGG 48
Db 71 GGGGGGGGGGGG--GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 114

```

[illegible][illegible]

Search completed: September 7, 2005, 12:59:57
Job time : 113.5 secs

This Page Blank (uspto)

XX PD 02-JAN-2003.
 XX PF 18-JUN-2002; 2002EP-00013517.
 XX PR 18-JUN-2001; 2001JP-00246789.
 XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 XX PA (AUSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;
 XX DR WPI; 2003-315783/31.
 XX DR N-PSDB; ADC87060.
 XX XX New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX PS Claim 2; SEQ ID NO 1514; 28pp; English.
 XX CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.
 XX SQ Sequence 920 AA;
 Query Match 77.3%; Score 230.5; DB 7; Length 920;
 Best Local Similarity 80.4%; Pred. No. 3.4e-14;
 Matches 41; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
 Qy 1 GGGGGGGGGGGGGGGGGGGGGGGRKRRRRGGGGGGGGGGGGGGGGGGGG 51
 Db 245 GGGGGGGGGGGGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGG 286
 RESULT 4
 ID ADG88556 standard; protein; 200 AA.
 XX AC ADG88556;
 XX DT 11-MAR-2004 (first entry)
 XX DE Poly Gly flexible linker.
 XX KW Human; angiogenesis; contactin associated protein 3; CASPR3;
 KW gene therapy.
 XX OS Unidentified.
 XX FH Key Location/Qualifiers
 FT Misc-difference 6..200
 FT /note= "Gly residues from position 6 to 200 may be
 FT present or absent"
 XX US2003176333-A1.
 XX PD 18-SEP-2003.
 XX PF 18-MAR-2002; 2002US-00100818.
 XX PR 18-MAR-2002; 2002US-00100818.
 XX PA (RIGE-) RIGEL PHARM INC.
 XX PI Lorens JB, Xu W, Bogenberger J;
 XX WPI; 2003-843828/78.

XX Methods for identifying compounds which regulate angiogenesis via
 PT modulation of contactin associated protein 3 (CASPR3) for the used in
 PT diagnosis and therapy of diseases related to angiogenesis.
 XX PS Disclosure; SEQ ID NO 14; 61pp; English.
 XX CC The present invention relates to a method of identifying a compound that
 CC regulates angiogenesis via the modulation of contactin associated protein
 CC 3 (CASPR3). The method involves contacting the compound with a CASPR3
 CC polypeptide and determining the functional effect of the compound upon
 CC the CASPR3 polypeptide. The CASPR3 nucleic acid and encoded protein is
 CC useful as a drug target for anti-angiogenic therapies. The invention is
 CC also useful in gene therapy. The present sequence is poly Gly flexible
 CC linker. This sequence is used to illustrate the method of the invention.
 XX SQ Sequence 200 AA;
 Query Match 76.7%; Score 228.5; DB 7; Length 200;
 Best Local Similarity 80.4%; Pred. No. 1.5e-14;
 Matches 41; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
 Qy 1 GGGGGGGGGGGGGGGGGGGGGGGRKRRRRGGGGGGGGGGGGGGGGGGGG 51
 Db 1 GGGGGGGGGGGGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGGGG 42
 RESULT 5
 ID ADJ84541 standard; protein; 200 AA.
 XX AC ADJ84541;
 XX DT 20-MAY-2004 (first entry)
 XX DE T2R G-protein coupled receptor related linker seq id 94.
 XX KW taste transduction; G-protein coupled receptor; T2R; taste signaling;
 KW cAMP level; cGMP level; IP3 level; Ca(2+) level; electrical activity;
 KW human; G-protein coupled receptor; receptor; linker.
 XX OS Synthetic.
 XX PN US2004038312-A1.
 XX PD 26-FEB-2004.
 XX PF 10-FEB-2003; 2003US-00364861.
 XX PR 10-SEP-1999; 99US-00393634.
 XX PA (ZUKE/) ZUKER C S.
 PA (ADLE/) ADLER J. E.
 PA (HOON/) HOON M.
 PA (RYBA/) RYBA N.
 PA (MUEL/) MUELLER K.
 XX PI Zuker CS, Adler JE, Hoon M, Ryba N, Mueller K;
 XX WPI; 2004-203221/19.
 XX PT Novel isolated taste transduction G-protein coupled receptor e.g., T2R
 PT useful for identifying compound that modulates taste signaling in taste
 XX cells.
 XX PS Disclosure; SEQ ID NO 94; 121pp; English.
 XX CC The invention describes an isolated taste transduction G-protein coupled
 CC receptor (I) e.g., T2R which is expressed in a taste cell, comprises
 CC greater than 60% sequence identity to a fully defined sequence of 335
 CC (S1), 333 (S2), 299 (S3), 310 (S4), 224 (S5), 77 (S6), 209 (S7), 266
 CC (S8), 300 (S9), 300 (S10), 155 (S11), 173 (S12), 94 (S13), 115 (S14), 68
 CC (S15), 126 (S16) and 180 (S17) amino acids as given in the specification.

CC (I) is useful for identifying a compound that modulates taste signaling
 CC in taste cells which involves contacting the compound with (I) and
 CC determining the functional effect of the compound (I). The functional
 CC effect is determined by measuring changes in intracellular cAMP, cGMP,
 CC IP3, or Ca²⁺. The functional effect is a chemical or physical effect.
 CC The functional effect is determined by measuring binding of the compound
 CC to an extracellular domain of (I). The functional effect is determined by
 CC measuring binding of radiolabeled GTP to (I). (I) is recombinant and is
 CC from rat, mouse, or human. (I) is expressed in a cell or cell membrane.
 CC The functional effect is measured by determining changes in the
 CC electrical activity of cells (e.g., eukaryotic cell) expressing (I). (I)
 CC is useful for identifying compound that modulates taste signaling in
 CC taste cells which involves contacting a compound with (I). The
 CC extracellular domain (II) of (I) is useful for identifying a compound
 CC that modulates taste signaling in taste cells which involves contacting
 CC the compound with the polynucleotide encoding (I) and determining the
 CC functional effect of the compound upon the extracellular domain. This is
 CC the amino acid sequence of a T2R G-protein coupled receptor related
 CC linker.

XX Sequence 200 AA;

Query Match 76.7%; Score 228.5; DB 8; Length 200;

Best Local Similarity 80.4%; Pred. No. 1.5e-14;

Matches 41; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

Qy 1 GGGGGGGGGGGGGGGGGGGGGRKRRRRGGGGGGGGGGGGGGGGGGGGG 51

Db 1 GGGGGGGGGGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGGGGG 42

RESULT 6

ADM96215

ID ADM96215 standard; protein; 200 AA.

XX

AC ADM96215;

XX

DT 01-JUL-2004 (first entry)

XX

DE Poly Gly flexible linker.

XX

KM G-protein coupled receptor 39; GPR 39; cancer; breast; ovary; prostate;
 KM brain; lung; ligand binding; aberrant receptor protein analysis;
 KM cytosstatic; receptor; poly Gly flexible linker.

XX

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Misc-difference 6. .200

FT /note= "Residues are independently optionally absent"

XX

PN US2004071708-A1.

XX

PD 15-APR-2004.

XX

XX 26-SEP-2002; 2002US-00255775.

XX

XX 26-SEP-2002; 2002US-00255775.

XX

XX (IMMU-) IMMUSOL INC.

XX

XX Claassen G, Li H, Barber J;

XX

XX WPI; 2004-373880/35.

XX

XX Identifying anticancer agents modulating G-protein coupled receptor
 PT proteins, comprises contacting GPR 39-specific binding agents to cancer
 PT cells and detecting anticancer activity to identify anticancer agent.

XX

PS Disclosure; SEQ ID NO 3; 33pp; English.

XX

XX The invention relates to a method of identifying anticancer agents that
 CC modulate G-protein coupled receptor (GPR) proteins, comprising contacting

CC GPR 39-specific binding agents to cancer cells and detecting anticancer
 CC activity. The invention also relates to a method of inhibiting cancer
 CC characteristics in cancer cells by down modulating GPR 39 protein
 CC activity to a level sufficient to inhibit the cancer characteristics of
 CC the cancer cells and an antibody that specifically recognises a GPR 39
 CC protein. The method further involves binding a population of different
 CC compositions to the GPR 39 protein to select GPR 39-specific binding
 CC agents. The cancer cells are chosen from breast, ovarian, prostate, brain
 CC and lung cancer cells. The method is useful for identifying anticancer
 CC agents that modulate GPR 39 protein and for inhibiting cancer
 CC characteristics in cancer cells. The antibody is useful for inhibiting
 CC receptor function e.g. by blocking ligand binding, which is used in
 CC therapeutic function, in various in vitro histochemical techniques for
 CC detecting GPR proteins and as a diagnostic tool for aberrant receptor
 CC protein analysis. This sequence represents a poly Gly flexible linker
 CC used in the method of the invention.

XX Sequence 200 AA;

Query Match 76.7%; Score 228.5; DB 8; Length 200;

Best Local Similarity 80.4%; Pred. No. 1.5e-14;

Matches 41; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

Qy 1 GGGGGGGGGGGGGGGGGGGGGRKRRRRGGGGGGGGGGGGGGGGGGGGG 51

Db 1 GGGGGGGGGGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGGGGG 42

RESULT 7

ABG71303

ID ABG71303 standard; protein; 201 AA.

XX

AC ABG71303;

XX

DT 17-DEC-2002 (first entry)

XX

DE PINPOINT poly-Gly linker protein.

XX

KM Protein position identification with nuclease tail; linker; PINPOINT;
 KM molecular beacon.

XX

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Misc-difference 1. .97

FT /note= "Optionally absent"

FT

FT Misc-difference 105. .201

FT /note= "Optionally absent"

XX

PN US6444421-B1.

XX

PD 03-SEP-2002.

XX

XX 02-APR-1998; 98US-00054281.

XX

XX 03-APR-1997; 97US-0080234P.

XX

XX 03-APR-1997; 97US-00825664.

XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

XX Chung JH;

XX

XX WPI; 2002-749290/81.

XX

XX Detecting whether a first molecule is in close proximity to a second
 PT molecule, or detecting target molecule, by using a molecular beacon
 PT comprising an oligonucleotide to which a fluorophore and a quencher are
 PT attached.

XX

PS Disclosure; Col 59-60; 41pp; English.

XX

XX The invention relates to detecting (M) whether a first molecule is in
 CC close proximity to a second molecule, or detecting a target molecule,

CC involves using a molecular beacon that comprises an oligonucleotide (ON)
CC to which a fluorophore and a quencher (FQ) are attached. The method is
CC named Protein Position Identification with Nuclease Tail (PINPOINT). The
CC method further comprises, whether a first molecule is in close proximity
CC to a second molecule involves attaching a molecular beacon to the first
CC molecule, where the molecular beacon comprises an (ON) to which is
CC attached an (FQ), attaching an endonuclease (EN) moiety to the second
CC molecule, and determining whether the first molecule is in close
CC proximity to the second molecule by detecting whether fluorescence is
CC emitted by the fluorophore, where fluorescence emission is indicative of
CC cleavage of the (ON) by the (EN) moiety, to cause separation of the (FQ).
CC The method may also comprise, detecting a target molecule, which involves
CC contacting the target molecule with a chimaeric fusion molecule that
CC comprises an (EN) molecule and a guide molecule that binds to the target
CC molecule, thus guiding the chimaeric fusion molecule to the target
CC molecule, contacting the chimaeric endonuclease with a molecular beacon
CC comprising an (ON) to which is attached a fluorophore and a quencher, and
CC detecting the presence of a fluorescent signal which results from
CC cleavage of the (ON) by the (EN), to allow separation of the quencher
CC from the fluorophore. The method is useful for detecting whether a first
CC molecule is in close proximity to a second molecule, or for detecting a
CC target molecule and is useful for detecting intermolecular interactions
CC in vivo, in situ and in vitro, such as enzymatic reactions, hormone-
CC ligand interactions, and drug or toxin interactions. The present sequence
CC is a proteinlinker used in the PINPOINT constructs of the invention
XX
XX Sequence 201 AA:
SQ

```
Query Match      76.7%   Score 228.5; DB 5; Length 201;
Best Local Similarity 80.4%; Pred. No. 1.5e-14;
Matches 41; Conservative 0; Mismatches 1; Indels 9;
```

RESULT 8	
ABU22369	
ID	ABU22369 standard; protein; 188 AA.
XX	
XX	ABU22369;
XX	
XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Protein encoded by Prokaryotic essential gene #7896.
XX	
KW	Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	
OS	Burkholderia mallei.
XX	
PN	WO200277183-A2.
PN	
PD	03-OCT-2002.
XX	
XX	
PF	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	
XX	(ELIT-) ELITRA PHARM INC.
PA	
XX	
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlssen KL, Zyskind JW;
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
XX	WPI; 2003-029926/02.
DR	N-PSDB; ACA26239.
DR	
XX	
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to

```

PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 50293; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 188 AA;
XX
XX Query Match 72.0%; Score 214.5; DB 6; Length 188;
XX Best Local Similarity 67.2%; Pred. No. 3.1e-13;
XX Matches 43; Conservative 2; Mismatches 6; Indels 13; Gaps 3;
XX
Qy 1 GGGGGGGGGGGGGG-----GGGGGGYGRKRRRRRGGG---GGGG-----GGGGGGGGG 47
Dy 115 GGGGGGGGGGGGGGDDGGYGGGGYGGGRDMMRGGGGGGRAGSGGGGGGGGGG 174
Qy 48 GGGG 51
Dy 175 GGGG 178
XX
XX RESULT 9
XX ABP29010
XX ID ABP29010 standard; protein; 56 AA.
XX AC ABP29010;
XX AD
XX AE
XX AF
XX AG
XX AH
XX AI
XX AJ
XX AK
XX AL
XX AM
XX AN
XX AO
XX AP
XX AQ
XX AR
XX AS
XX AT
XX AU
XX AV
XX AW
XX AX
XX AY
XX AZ
XX BA
XX BB
XX BC
XX BD
XX BE
XX BF
XX BG
XX BH
XX BI
XX BJ
XX BK
XX BL
XX BM
XX BN
XX BO
XX BP
XX BQ
XX BR
XX BS
XX BT
XX BU
XX BV
XX BW
XX BX
XX BY
XX BZ
XX CA
XX CB
XX CC
XX CD
XX CE
XX CF
XX CG
XX CH
XX CI
XX CJ
XX CK
XX CL
XX CM
XX CN
XX CO
XX CP
XX CQ
XX CR
XX CS
XX CT
XX CU
XX CV
XX CW
XX CX
XX CY
XX CZ
XX DA
XX DB
XX DC
XX DD
XX DE
XX DF
XX DG
XX DH
XX DI
XX DJ
XX DK
XX DL
XX DM
XX DN
XX DO
XX DP
XX DQ
XX DR
XX DS
XX DT
XX DU
XX DV
XX DW
XX DX
XX DY
XX DZ
XX EA
XX EB
XX EC
XX ED
XX EE
XX EF
XX EG
XX EH
XX EI
XX EJ
XX EK
XX EL
XX EM
XX EN
XX EO
XX EP
XX EQ
XX ER
XX ES
XX ET
XX EU
XX EV
XX EW
XX EX
XX EY
XX EZ
XX FA
XX FB
XX FC
XX FD
XX FE
XX FF
XX FG
XX FH
XX FI
XX FJ
XX FK
XX FL
XX FM
XX FN
XX FO
XX FP
XX FQ
XX FR
XX FS
XX FT
XX FU
XX FV
XX FW
XX FX
XX FY
XX FZ
XX GA
XX GB
XX GC
XX GD
XX GE
XX GF
XX GG
XX GH
XX GI
XX GJ
XX GK
XX GL
XX GM
XX GN
XX GO
XX GP
XX GQ
XX GR
XX GS
XX GT
XX GU
XX GV
XX GW
XX GX
XX GY
XX GZ
XX HA
XX HB
XX HC
XX HD
XX HE
XX HF
XX HG
XX HH
XX HI
XX HJ
XX HK
XX HL
XX HM
XX HN
XX HO
XX HP
XX HQ
XX HR
XX HS
XX HT
XX HU
XX HV
XX HW
XX HX
XX HY
XX HZ
XX IA
XX IB
XX IC
XX ID
XX IE
XX IF
XX IG
XX IH
XX II
XX IJ
XX IK
XX IL
XX IM
XX IN
XX IO
XX IP
XX IQ
XX IR
XX IS
XX IT
XX IU
XX IV
XX IW
XX IX
XX IY
XX IZ
XX JA
XX JB
XX JC
XX JD
XX JE
XX JF
XX JG
XX JH
XX JI
XX JJ
XX JK
XX JL
XX JM
XX JN
XX JO
XX JP
XX JQ
XX JR
XX JS
XX JT
XX JU
XX JV
XX JW
XX JX
XX JY
XX JZ
XX KA
XX KB
XX KC
XX KD
XX KE
XX KF
XX KG
XX KH
XX KI
XX KJ
XX KK
XX KL
XX KM
XX KN
XX KO
XX KP
XX KQ
XX KR
XX KS
XX KT
XX KU
XX KV
XX KW
XX KX
XX KY
XX KZ
XX LA
XX LB
XX LC
XX LD
XX LE
XX LF
XX LG
XX LH
XX LI
XX LJ
XX LK
XX LL
XX LM
XX LN
XX LO
XX LP
XX LQ
XX LR
XX LS
XX LT
XX LU
XX LV
XX LW
XX LX
XX LY
XX LZ
XX MA
XX MB
XX MC
XX MD
XX ME
XX MF
XX MG
XX MH
XX MI
XX MJ
XX MK
XX ML
XX MM
XX MN
XX MO
XX MP
XX MQ
XX MR
XX MS
XX MT
XX MU
XX MV
XX MW
XX MX
XX MY
XX MZ
XX NA
XX NB
XX NC
XX ND
XX NE
XX NF
XX NG
XX NH
XX NI
XX NJ
XX NK
XX NL
XX NM
XX NO
XX NP
XX NQ
XX NR
XX NS
XX NT
XX NU
XX NV
XX NW
XX NX
XX NY
XX NZ
XX OA
XX OB
XX OC
XX OD
XX OE
XX OF
XX OG
XX OH
XX OI
XX OJ
XX OK
XX OL
XX OM
XX ON
XX OO
XX OP
XX OQ
XX OR
XX OS
XX OT
XX OU
XX OV
XX OW
XX OX
XX OY
XX OZ
XX PA
XX PB
XX PC
XX PD
XX PE
XX PF
XX PG
XX PH
XX PI
XX PJ
XX PK
XX PL
XX PM
XX PN
XX PO
XX PP
XX PQ
XX PR
XX PS
XX PT
XX PU
XX PV
XX PW
XX PX
XX PY
XX PZ
XX QA
XX QB
XX QC
XX QD
XX QE
XX QF
XX QG
XX QH
XX QI
XX QJ
XX QK
XX QL
XX QM
XX QN
XX QO
XX QP
XX QQ
XX QR
XX QS
XX QT
XX QU
XX QV
XX QW
XX QX
XX QY
XX QZ
XX RA
XX RB
XX RC
XX RD
XX RE
XX RF
XX RG
XX RH
XX RI
XX RJ
XX RK
XX RL
XX RM
XX RN
XX RO
XX RP
XX RQ
XX RR
XX RS
XX RT
XX RU
XX RV
XX RW
XX RX
XX RY
XX RZ
XX SA
XX SB
XX SC
XX SD
XX SE
XX SF
XX SG
XX SH
XX SI
XX SJ
XX SK
XX SL
XX SM
XX SN
XX SO
XX SP
XX SQ
XX SR
XX SS
XX ST
XX SU
XX SV
XX SW
XX SX
XX SY
XX SZ
XX TA
XX TB
XX TC
XX TD
XX TE
XX TF
XX TG
XX TH
XX TI
XX TJ
XX TK
XX TL
XX TM
XX TN
XX TO
XX TP
XX TQ
XX TR
XX TS
XX TT
XX TU
XX TV
XX TW
XX TX
XX TY
XX TZ
XX UA
XX UB
XX UC
XX UD
XX UE
XX UF
XX UG
XX UH
XX UI
XX UJ
XX UK
XX UL
XX UM
XX UN
XX UO
XX UP
XX UQ
XX UR
XX US
XX UT
XX UY
XX UZ
XX VA
XX VB
XX VC
XX VD
XX VE
XX VF
XX VG
XX VH
XX VI
XX VJ
XX VK
XX VL
XX VM
XX VN
XX VO
XX VP
XX VQ
XX VR
XX VS
XX VT
XX VY
XX VZ
XX WA
XX WB
XX WC
XX WD
XX WE
XX WF
XX WG
XX WH
XX WI
XX WJ
XX WK
XX WL
XX WM
XX WN
XX WO
XX WP
XX WQ
XX WR
XX WS
XX WT
XX WY
XX WZ
XX XA
XX XB
XX XC
XX XD
XX XE
XX XF
XX XG
XX XH
XX XI
XX XJ
XX XK
XX XL
XX XM
XX XN
XX XO
XX XP
XX XQ
XX XR
XX XS
XX XT
XX XU
XX XV
XX XW
XX XX
XX XY
XX XZ
XX YA
XX YB
XX YC
XX YD
XX YE
XX YF
XX YG
XX YH
XX YI
XX YJ
XX YK
XX YL
XX YM
XX YN
XX YO
XX YP
XX YQ
XX YR
XX YS
XX YT
XX YU
XX YV
XX YW
XX YX
XX YY
XX YZ
XX ZA
XX ZB
XX ZC
XX ZD
XX ZE
XX ZF
XX ZG
XX ZH
XX ZI
XX ZJ
XX ZK
XX ZL
XX ZM
XX ZN
XX ZO
XX ZP
XX ZQ
XX ZR
XX ZS
XX ZT
XX ZU
XX ZV
XX ZW
XX ZX
XX ZY
XX ZZ

```

```

Query Match          72.0%; Score 214.5; DB 6; Length 188;
Best Local Similarity 67.2%; Pred. No. 3.le-13;
Matches 43; Conservative 2; Mismatches 6; Indels 13; Gaps 3;

Qy      1  GGGGGGGGGGGG-----GGGGGGYGRKKRRRGGG---GGG-----GGGGGGGGG 47
         |||||
Db      115 GGGGGGGGGGGGDGGYGGGGGYGGGRDMERGGGGGRASGGGGAGARSGGGGGGGG 174
           ||||| : ||| ||| |||||
Qy      48  GGGG 51
         |||||
Db      175 GGGG 178
           |||||

RESULT 9
ABP29010
ID   ABP29010 standard; protein; 56 AA.
XX
AC   ABP29010;
XX
DT   02-JUL-2002 (first entry)
XX
DE   Streptococcus polypeptide SEQ ID NO 7196.
XX
KW   Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW   group A streptococcus; Streptococcus pyogenes; antibacterial;
KW   antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS   Streptococcus agalactiae.
XX
PN   WO200234771-A2.
XX
PD   02-MAY-2002.
XX
PF   29-OCT-2001; 2001WO-GB004789.
XX

```



```

PN WO200171042-A2.
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US009231.
PF
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
PR
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
XX N-PSDB; ABL08832.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
PT
XX
XX Disclosure; SEQ ID NO 20979; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 344 AA;
SQ
XX
XX Query Match 66.8%; Score 199; DB 4; Length 344;
XX Best Local Similarity 65.1%; Pred. No. 1.5e-11;
XX Matches 41; Conservative 1; Mismatches 5; Indels 16; Gaps 3;
XX
XX 1 GGGGGGGGGG-----GGGGGGGGGGYGRKRRRQRRGGG-----GGGGGGGGGGGG 48
Qy |||||
Db 12 GGGGGGGGGGPRGGGGGGGGGGFG-----GGRGGGGGDRGGGGGGGGGGGG 67
XX
XX 49 GGG 51
Qy |||
Db 68 GGG 70
XX
XX
XX RESULT 12
XX ABB60010
XX ID ABB60010 standard; protein; 399 AA.
XX AC ABB60010;
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 6822.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
FN
XX
XX 27-SEP-2001.
FD
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
PR
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
XX N-PSDB; ABL08832.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
PT
XX
XX Disclosure; SEQ ID NO 6822; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 399 AA;
SQ
XX
XX Query Match 66.6%; Score 198.5; DB 4; Length 399;
XX Best Local Similarity 67.2%; Pred. No. 1.9e-11;
XX Matches 41; Conservative 0; Mismatches 3; Indels 17; Gaps 3;
XX
XX 1 GGGGGGGGGG-----GGGGGGGGGGYGRKRRRQRRGGGGGG-----GGGGGGGGGGGG 48
Qy |||||
Db 215 GGGGGGGGGGFRGGGGGGGGGGGGGGGGGGFD-----RGGGGGGGGYDRGGGGGGGGGG 269
XX
XX 49 G 49
Qy |
Db 270 G 270
XX
XX
XX RESULT 13
XX ABO59421
XX ID ABO59421 standard; protein; 208 AA.
XX AC ABO59421;
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Human genome derived single exon protein #5655.
DE
XX
XX Human; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
KW
XX
XX Homo sapiens.
OS
XX
XX US2003194704-A1.
FN
XX
XX 16-OCT-2003.
PD
XX
XX 03-APR-2002; 2002US-00029386.
PF
XX
XX 03-APR-2002; 2002US-00029386.
PR
XX
XX (PENN/) PENN S G.
PA
XX (RANK/) RANK D R.
PA
XX (HANZ/) HANZEL D K.
PA
XX
XX Penn SG, Rank DR, Hanzel DK;
PI
XX
XX WPI; 2004-119264/12.
DR
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
PT
```


PI Zoghbi HY, Van Den Veyver IB, Amir R, Francke U;
XX WPI; 2004-256068/24.
DR

PT Screening human for Rett syndrome comprises detecting mutation in nucleic
XX acid sequence encoding methyl-CpG-binding protein 2 (MECP2).

PS Disclosure; SEQ ID NO 112; 125pp; English.

XX The invention relates to a method of screening a human for Rett syndrome
CC comprising detecting a mutation in a nucleic acid sequence encoding
CC methyl-CpG-binding protein 2 (MECP2). The method is useful for screening
CC a human for Rett syndrome. The method is useful for screening
CC neurodevelopmental diseases such as Rett syndrome, autism, non-syndromic
CC mental retardation, idiopathic neonatal encephalopathy, idiopathic
CC infantile spasms, idiopathic cerebral palsy, Angelman syndrome and
CC schizophrenia. The present sequence represents the amino acid sequence of
CC a methyl-CpG-binding protein.

XX SQ Sequence 345 AA;

Query Match 65.4%; Score 195; DB 8; Length 345;
Best Local Similarity 34.6%; Pred. NO. 3.6e-11;
Matches 44; Conservative 2; Mismatches 5; Indels 76; Gaps 2;

Qy 1 GGGGGGGGGGGGGG--GGGGYGR----- 23
Db 197 GGGGGGGGGGGGGGGGGGGGRRVGGGGGGRVRAAERGGGRLLVKMPFAGGGAPASPPAPTLP 256
Qy 24 -----KRRRRRRGGGGGGGGGGGG 44
Db 257 PSAHPPPTAPPATHGQGLGGGVKRPGRKRAEADSRVPKRGRKPGGGGGGGGGGG 316

Qy 45 GGGGGGG 51
Db 317 GGGGVGG 323

Search completed: September 7, 2005, 12:56:08
Job time : 66.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 12:48:54 ; Search time 29.5 Seconds
(without alignments)
129.054 Million cell updates

Title: US-09-910-432-20

Perfect score: 298

Sequence: 1 GCGGGGCGGGCGGGGG...GCGGGGCGGGCGGGGG 51

Scoring table: BLOSUM62

Learning curve: 2500000
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs. 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database : Issued_Parents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pcp.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pcp.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pcp.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pcp.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pcp.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	228.5	76.7	200	4	US-09-989-981A-13	Sequence 13, Appl
2	228.5	76.7	201	3	US-09-052-995-1	Sequence 1, Appl
3	228.5	76.7	201	3	US-09-053-003-40	Sequence 40, Appl
4	228.5	76.7	201	4	US-09-054-281-22	Sequence 22, Appl
5	228.5	76.7	201	4	US-09-478-948-6	Sequence 6, Appl
6	228.5	76.7	201	4	US-09-618-094-40	Sequence 40, Appl
7	228.5	76.7	201	4	US-09-754-947-5	Sequence 5, Appl
8	195	65.4	345	4	US-09-657-013-112	Sequence 112, App
9	194	65.1	162	3	US-09-575-574-4	Sequence 4, Appl
10	192.5	64.6	426	4	US-09-319-039-324	Sequence 324, App
11	189.5	63.6	526	4	US-09-538-092-1080	Sequence 1080, Ap
12	183.5	61.6	112	4	US-09-543-681A-8319	Sequence 8319, Ap
13	183	61.4	268	2	US-08-835-099A-9	Sequence 9, Appl
14	183	61.4	268	3	US-09-157-349-9	Sequence 9, Appl
15	181.5	60.9	60	4	US-09-832-297A-12	Sequence 12, Appl
16	181.5	60.9	100	4	US-09-411-067C-4	Sequence 4, Appl
17	181.5	60.9	738	3	US-08-864-038A-3	Sequence 3, Appl
18	178.5	59.9	68	4	US-09-107-433-3	Sequence 4787, Ap
19	178.5	59.9	111	4	US-09-543-681A-8322	Sequence 8322, Ap
20	178	59.7	266	3	US-09-032-523-7	Sequence 7, Appl
21	178	59.7	266	4	US-09-802-633-7	Sequence 7, Appl
22	177.5	59.6	179	4	US-09-270-767-57837	Sequence 57837, A
23	176.5	59.2	40	3	US-08-815-190A-17	Sequence 17, Appl
24	176.5	59.2	40	4	US-09-333-213-1	Sequence 1, Appl
25	173	58.1	1958	1	US-07-945-283-2	Sequence 2, Appl
26	171.5	57.6	334	4	US-09-949-016-11006	Sequence 11006, A
27	171	57.4	204	4	US-09-107-433-4573	Sequence 4573, Ap

28	167.5	56.2	420	2	US-08-845-998-8	Sequence 8, Appli
29	167.5	56.2	420	3	US-08-206-537-8	Sequence 8, Appli
30	167.5	56.2	420	3	US-08-430-854-8	Sequence 8, Appli
31	167.5	56.2	637	4	US-08-949-016-8152	Sequence 8152, App
32	166	55.7	91	4	US-09-270-767-39789	Sequence 39789, A
33	166	55.7	91	4	US-09-270-767-55097	Sequence 55006, A
34	165	55.4	316	4	US-09-538-092-997	Sequence 997, App
35	164	55.0	643	4	US-09-538-092-844	Sequence 844, App
36	163.5	54.9	103	4	US-09-107-433-4326	Sequence 4326, App
37	162	54.4	647	2	US-08-770-761A-8	Sequence 8, Appli
38	162	54.4	705	2	US-08-770-761A-7	Sequence 7, Appli
39	159.5	53.5	714	2	US-08-900-114-3	Sequence 3, Appli
40	159.5	53.5	714	3	US-08-241-333-3	Sequence 3, Appli
41	158	53.0	300	3	US-08-765-856-2	Sequence 2, Appli
42	158	53.0	300	3	US-08-935-009A-2	Sequence 2, Appli
43	158	53.0	302	3	US-08-765-856-4	Sequence 4, Appli
44	158	53.0	302	3	US-08-935-009A-1	Sequence 4, Appli
45	158	53.0	464	4	US-09-252-931A-24883	Sequence 24883, A

ALIGNMENTS

RESULT 1

```

US-09-989-981A-13
; Sequence 13, Application US/0998981A
; Patent No. 6821750
; GENERAL INFORMATION:
; APPLICANT: Hobbs, Helen H.
; APPLICANT: Shan, Bei
; APPLICANT: Barnes, Robert
; APPLICANT: Tian, Hui
; APPLICANT: Tullarik Inc.
; TITLE OF INVENTION: The University of Texas System
; OF INVENTION: ABCG5 and ABCG8: Compositions and Methods of Use
; FILE REFERENCE: 018781-007320US
; CURRENT APPLICATION NUMBER: US/09/989, 981A
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 60/252,235
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/253,645
; PRIOR FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 200

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:flexible linker
;
; FEATURE:
; OTHER INFORMATION: Gly at positions 6-200 may be present or absent
;
; NAME/KEY: MOD RES
; LOCATION: (6)..(200)
; OTHER INFORMATION:
US-09-989-981A-13

```

Query Match	Score	DB 4	Length
Best Local Similarity	76.7%	228.5	200
Matches	80.4%	Pred. No. 2.3e-15	
Conservative	41	0	1
Mismatches	0	1	9
Indels	0	1	9
Gaps	0	1	9

[illegible]

RESULT 2

US-09-052-995-1
; Sequence 1, Application US/09052995
; Patent No. 6183956
; GENERAL INFORMATION:
; APPLICANT: Sivaraja, Mohanram
; APPLICANT: Strulovici, Berta
; APPLICANT: Flores, Osvaldo A.

[illegible]

```

SOFTWARE: Patent Release #1.0, Version 1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,281
FILING DATE: 02-APR-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/080,234
FILING DATE: 03-APR-1997
PRIOR APPLICATION NUMBER: US 08/825,622
FILING DATE: 03-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 015280-2952000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 22:

```



```
; FILE REFERENCE: 014907-003310US
; CURRENT APPLICATION NUMBER: US/09/754,947
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,901
; PRIOR FILING DATE: 2000-01-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: flexible linker
; NAME/KEY: MOD RES
; LOCATION: (1)..(97)
; OTHER INFORMATION: Gly at positions 1-97 may be present or absent
; NAME/KEY: MOD RES
; LOCATION: (105)..(201)
; OTHER INFORMATION: Gly at positions 105-201 may be present or absent
US-09-754-947-5

Query Match          76.7%; Score 228.5; DB 4; Length 201;
Best Local Similarity 80.4%; Pred. No. 2.3e-15;
Matches 41; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGVRKKRRRRRRGGGGGGGGGGGGGGGGGGGGGG 51
Db 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 42

RESULT 8
US-09-657-013-112
; Sequence 112, Application US/09657013
; Patent No. 6709817
; GENERAL INFORMATION:
; APPLICANT: Zoghbi, Huda Y.
; APPLICANT: Van den Veyver, Ignatia B
; APPLICANT: Amir, Ruthie
; APPLICANT: Francke, Uta
; TITLE OF INVENTION: Methods of Identifying Mutations in a Methyl-CPG-Binding Domain
; FILE REFERENCE: HO-P01893US1/0905371
; CURRENT APPLICATION NUMBER: US/09/657,013
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: US 60/152,778
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 112
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Chicken
US-09-657-013-112

Query Match          65.4%; Score 195; DB 4; Length 345;
Best Local Similarity 34.6%; Pred. No. 6e-12;
Matches 44; Conservative 2; Mismatches 5; Indels 76; Gaps 2;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 23
Db 197 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 256
Qy 24 -----KKRRRRRRGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 44
Db 257 PSAAHPPTTAPPATHGGGLGGGVKRPGRKKAADSRVPKRRKPGGGGGGGGGGG 316
Qy 45 GGGGGGGG 51
Db 317 GGGGGVGG 323

RESULT 9
US-09-575-574-4
```

```
; Sequence 4, Application US/09575574
; Patent No. 6376750
; GENERAL INFORMATION:
; APPLICANT: Yu, Su-May
; APPLICANT: Chao, Yu-Chan
; TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER
; FILE REFERENCE: 08919-047001
; CURRENT APPLICATION NUMBER: US/09/575,574
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-575-574-4

Query Match          65.1%; Score 194; DB 3; Length 162;
Best Local Similarity 65.0%; Pred. No. 4e-12;
Matches 39; Conservative 0; Mismatches 5; Indels 16; Gaps 2;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 50
Db 90 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 143

RESULT 10
US-09-919-039-324
; Sequence 324, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; PRIOR FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 324
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 181344CD1
US-09-919-039-324

Query Match          64.6%; Score 192.5; DB 4; Length 462;
Best Local Similarity 48.2%; Pred. No. 1.3e-11;
Matches 41; Conservative 3; Mismatches 6; Indels 35; Gaps 3;

Qy 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 38
Db 166 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225
Qy 39 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 50
Db 226 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 250

RESULT 11
US-09-538-092-1080
; Sequence 1080, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
```

```
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 1080
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P35637
; US-09-538-092-1080

Query Match      63.6%; Score 189.5; DB 4; Length 526;
Best Local Similarity 48.2%; Pred. No. 2.9e-11;
Matches 41; Conservative 2; Mismatches 7; Indels 35; Gaps 3;

Qy 1 GGGGGGGGGG-----GGG-----GGGGGGYGRKKRRRGGGGGG 38
    |||||
Db 166 GGGGGGGGGYQDQSSMSGGSGGYGNQDQSGGGSGGYGQDGRGRGGGGGG 225
    |||||
Qy 39 GGGGG-----GGGGG 50
    |||||
Db 226 GGGGGGYNRRSGGYEPRGGGRG 250
    |||||

RESULT 12
US-09-543-681A-8319
; Sequence 8319, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 8319
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-8319

Query Match      61.6%; Score 183.5; DB 4; Length 112;
Best Local Similarity 69.2%; Pred. No. 3e-11;
Matches 36; Conservative 1; Mismatches 4; Indels 11; Gaps 2;

Qy 1 GGGGGGGGGGGG--GGGGYGRKKRRRGGGGGGGGGGGGGGGGGGGG 50
    |||||
Db 5 GGGGGGGGGGGGKKKGW-----RGRGRGGGGGGGGGGGG 47
    |||||

RESULT 13
US-08-835-099A-9
; Sequence 9, Application US/08835099A
; Patent No. 5874277
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasuehi
; APPLICANT: NISHI, Kazuhiro
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston

; PRIOR APPLICATION NUMBER: 08/835,099
; FILING DATE:
```

```
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,099A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 083649/1996
; FILING DATE: 05-APR-1996
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-835-099A-9

Query Match      61.4%; Score 183; DB 2; Length 268;
Best Local Similarity 73.9%; Pred. No. 6.9e-11;
Matches 34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 6 GGGGGGGGGGGYGRKKRRRGGGGGGGGGGGGGGGGGGGGGGGG 51
    |||||
Db 10 GGGGGGGGGGGGLGNVLGSLGAGGGGGGGGGGGGGGGGGGG 55
    |||||

RESULT 14
US-09-157-349-9
; Sequence 9, Application US/09157349
; Patent No. 6068990
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasuehi
; APPLICANT: NISHI, Kazuhiro
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
```


This Page Blank (uspto)

Result No.	Score	Query Match	Length	DB	ID	Description
1	221	74.2	221	2	T04592	glycine-rich cell
2	205	68.8	155	1	KNR2G1	glycine-rich cell
3	204.5	68.6	404	2	S54729	RNA-binding protein
4	203.5	68.3	165	2	T03583	glycine-rich RNA-b
5	203	68.1	136	2	T29282	hypothetical prote
6	202	67.8	396	2	T49109	glycine-rich prote
7	198	66.4	207	2	T07381	glycine-rich prote
8	197	66.1	440	2	S71795	transcription fact
9	196.5	65.9	285	2	T31503	hypothetical prote
10	196	65.8	255	2	B84777	hypothetical prote
11	194.5	65.3	168	1	S12312	glycine-rich RNA-b
12	194.5	65.3	291	1	S31415	glycine-rich prote
13	193.5	64.9	385	2	T20410	hypothetical prote
14	193	64.8	142	2	S12311	glycine-rich RNA-b
15	192	64.4	1226	2	T24045	hypothetical prote
16	191.5	64.3	106	2	F84797	hypothetical prote
17	191	64.1	167	2	S71779	glycine-rich RNA-b
18	190.5	63.9	384	1	A26099	glycine-rich cell
19	189.5	63.6	162	2	T10479	glycine-rich RNA-b
20	189.5	63.6	462	4	S33798	FUS/CHOP mutant fu
21	189.5	63.6	526	1	S33799	RNA-binding protein
22	189	63.4	252	1	S01821	glycine-rich prote
23	189	63.4	528	2	G02127	fus-like protein -
24	188	63.1	220	2	A44805	eggshell protein p
25	187	62.8	148	2	S41772	glycine-rich RNA-b
26	187	62.8	155	2	S20846	glycine-rich prote
27	187	62.8	214	1	KNR728	glycine-rich prote
28	187	62.6	239	2	S49193	GCR 101 protein -
29	186.5	62.6	157	1	S14957	glycine-rich prote

R;McMurray, A.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z21041
A:Accession: T31503
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-285 <WIL>
A:Cross-references: UNIPROT:Q9U2U0; EMBL:AL117204; PIDN:CA855137.1; CESP:Y116A8C.35
A:Experimental source: clone Y116A8C
C:Genetics:
A:Gene: CESP:Y116A8C.35
A:Introns: 31/3; 143/2

Query Match 65.9%; Score 196.5; DB 2; Length 285;
Best Local Similarity 73.1%; Pred. No. 2.2e-09;
Matches 38; Conservative 1; Mismatches 12; Indels 1; Gaps 1;

Qy 1 GGGGGGGGGGGGGGGGGY-GRKKRQRERGGGGGGGGGGGGGGGGGG 51
Db 213 GGG 264

RESULT 10
B84777
hypothetical protein At2g36120 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84777
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, I.;
euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84777
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <STO>
A:Cross-references: UNIPROT:Q9SIH2; GB:AE002093; NID:g4678224; PIDN:AAD26969.1; GSPDB:G
C:Genetics:
A:Gene: At2g36120
A:Map position: 2
C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology

Query Match 65.8%; Score 196; DB 2; Length 255;
Best Local Similarity 68.4%; Pred. No. 2.3e-09;
Matches 39; Conservative 1; Mismatches 11; Indels 6; Gaps 2;

Qy 1 GGGGGGGGGGGGGGGGGG--GGYGRKRRRRRRGGGGG-----GGGGGGGGGGGGGG 51
Db 125 GGGGGGGAGGGGGGGGGAHGGYGGGCGAGAGGGYGGGGAGGGGGGGGGGGGGGG 181

RESULT 11
S12312
glycine-rich RNA-binding protein (clone S2) - sorghum
C:Species: Sorghum bicolor (sorghum)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S12312
R;Cretin, C.; Puigdomenech, P.
Plant Mol. Biol. 15, 783-785, 1990
A:Title: Glycine-rich RNA-binding proteins from Sorghum vulgare.
A:Reference number: S12311; MUID:91346715; PMID:1715211
A:Accession: S12312
A:Molecule type: mRNA
A:Residues: 1-168 <CRE>
A:Cross-references: UNIPROT:Q99070; EMBL:X57662; NID:g21624; PIDN:CAA40862.1; PID:g21622
A:Note: in the authors' translation two additional Gly are shown after 110-Gly
C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
C:Keywords: GTP binding
F;9-76/Domain: ribonucleoprotein repeat homology <RRM1>

Query Match 65.3%; Score 194.5; DB 1; Length 168;

```
Best Local Similarity 63.1%; Pred. No. 2.2e-09;
Matches 41; Conservative 0; Mismatches 5; Indels 19; Gaps 3;

Qy 1 GGGGGGGGGGGG-----GGGGGGYGRKKRRQRRRGGGGGGGGGGGGGGG-----47
Db 89 GGGGGGGGGGGGYYGREGGGYGGGGGGY-----GRRREGGGGGYGGGGGGYGGR 143
Qy 48 -GGGG 51
Db 144 EGGGG 148

RESULT 12
S31415
glycine-rich protein GRP22 - rape
C/Species: Brassica napus (rape)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: S31415
R/Bergerson, D.; Boivin, R.; Baszczynski, C.L.; Bellemare, G.
submitted to the EMBL Data Library, August 1992
A/Description: Characterization and expression of a gene family encoding glycine-rich pr
A/Reference number: S31415
A/Accession: S31415
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-291 <BER>
A/Cross-references: UNIPROT:Q39337; EMBL:Z15045; NID:G17820; PIDN:CAA78762.1; PID:G17821
C/Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 65.3%; Score 194.5; DB 1; Length 291;
Best Local Similarity 65.0%; Pred. No. 3.3e-09;
Matches 39; Conservative 0; Mismatches 12; Indels 9; Gaps 2;

Qy 1 GGGGGGGGGGGGGGGG-----GGYGRKKRRQRRRGGGG-----GGGGGGGGGGGGGGG 51
Db 207 GGGGGGGGGGGGGGGGAGAGHGGYGAGGAGGGYGGGGGGGGGGGGGGGGGGGGGGG 266

RESULT 13
T20410
hypothetical protein E02A10.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T20410
R/Thomas, K.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19271
A/Accession: T20410
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-385 <WIL>
A/Cross-references: UNIPROT:Q93424; EMBL:Z81053; PIDN:CAB02877.1; GSPDB:GN00023; CESP:E0
A/Experimental source: clone E02A10
C/Genetics:
A/Gene: CESP:E02A10.2
A/Map position: 5
A/Introns: 32/1; 72/1; 85/1; 122/1; 133/1; 220/3

Query Match 64.9%; Score 193.5; DB 2; Length 385;
Best Local Similarity 72.5%; Pred. No. 4.8e-09;
Matches 37; Conservative 0; Mismatches 5; Indels 9; Gaps 1;

Qy 1 GGGGGGGGGGGGGGGGGGGGGYGRKKRRQRRRGGGGGGGGGGGGGGGGGGGGGGG 51
Db 136 GGGGGGGGGGGGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGGGGG 177

RESULT 14
S12311
glycine-rich RNA-binding protein (clone S1) - sorghum (fragment)
C/Species: Sorghum bicolor (sorghum)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C/Accession: S12311
```

```
R/Cretin, C.; Puigdomenech, P.
Plant Mol. Biol. 15, 783-785, 1990
A/Title: Glycine-rich RNA-binding proteins from Sorghum vulgare.
A/Reference number: S12311; MUID:91346715; PMID:1715211
A/Accession: S12311
A/Molecule type: mRNA
A/Residues: 1-142 <CRE>
A/Cross-references: UNIPROT:Q99069; EMBL:X57663; NID:Q21622; PIDN:CAA40863.1; PID:G21623
A/Note: the authors translated the codon ACG for residue 37 as Ser and CGC for residue 51
A/Note: the authors did not translate the codons for residues 92 and 93
C/Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
C/Keywords: RNA binding
F/1-55/Domain: ribonucleoprotein repeat homology (fragment) <RRM1>

Query Match 64.8%; Score 193; DB 2; Length 142;
Best Local Similarity 70.9%; Pred. No. 2.6e-09;
Matches 39; Conservative 0; Mismatches 6; Indels 10; Gaps 2;

Qy 1 GGGGGGGGGGGGGGGGGGGGGYGRKKRRQRRRGGGGGGGGGGGGGGG-----GGGG 51
Db 67 GGRGGGGGGGGYGGGGGGGGY-----RRDGGGGGGYGGGGGGYGGGGGGGGG 115

RESULT 15
T24045
hypothetical protein R08B4.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T24045
R/White, S.
submitted to the EMBL Data Library, November 1995
A/Reference number: Z19834
A/Accession: T24045
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-1226 <WIL>
A/Cross-references: UNIPROT:Q21835; EMBL:Z68008; PIDN:CAA92000.1; GSPDB:GN00028; CESP:R08
A/Experimental source: clone R08B4
C/Genetics:
A/Gene: CESP:R08B4.1
A/Map position: X
A/Introns: 53/2; 113/1; 152/3; 204/3; 272/1; 354/1; 389/3; 714/3; 839/3; 877/1; 960/3; 1

Query Match 64.4%; Score 192; DB 2; Length 1226;
Best Local Similarity 66.7%; Pred. No. 1.4e-08;
Matches 34; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGGGYGRKKRRQRRRGGGGGGGGGGGGGGGGGGGGGGG 51
Db 857 GCGNGGGGGGGGGGGGGGGGGGGGSGGSGGSGGSGGSGGSGGSGGSGGSGGNGGGG 907

Search completed: September 7, 2005, 13:00:51
Job time : 24 secs
```


Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	234	78.5	321	2	Q69XV3	Q69XV3 oryza sativa
2	221	74.2	221	2	Q65514	Q65514 arabidopsis
3	214.5	72.0	339	2	Q62FE6	Q62FE6 burkholderia
4	209.5	70.3	199	2	Q85215	Q85215 oryza sativa
5	205	68.8	165	1	GRP1_ORYSA	P25074 oryza sativa
6	204.5	68.6	175	2	Q9LGN6	Q9LGN6 arabidopsis
7	203.5	68.3	165	2	Q24184	Q24184 oryza sativa
8	203.5	68.3	551	2	Q275HG8	Q275HG8 oryza sativa
9	203	68.1	136	2	Q18444	Q18444 caenorhabditis
10	202	67.8	396	2	Q65450	Q65450 arabidopsis
11	200.5	67.3	535	2	Q942Q2	Q942Q2 oryza sativa
12	200.5	67.3	698	2	Q8QXK8	Q8QXK8 ectocarpus
13	199	66.8	165	2	Q95UX2	Q95UX2 drosophila
14	199	66.8	344	1	FBRL_DROME	Q9W1V3 drosophila
15	199	66.8	345	1	FBRL_DROBR	Q811F4 drosophila
16	198.5	66.6	399	5	CAZ_DROME	Q27294 drosophila
17	198	66.4	207	2	Q43522	Q43522 lycopersicon
18	197	66.1	155	2	Q9XEL4	Q9XEL4 picea glauca
19	197	66.1	239	2	Q69T79	Q69T79 oryza sativa
20	197	66.1	440	1	FXGA_CHICK	Q98937 gallus gallus
21	197	66.1	592	2	Q90W75	Q90W75 oncorhynchus
22	196.5	65.9	161	2	Q22385	Q22385 oryza sativa
23	196.5	65.9	285	2	Q9U2U0	Q9U2U0 caenorhabditis
24	196.5	65.9	422	2	Q6Z142	Q6Z142 oryza sativa
25	196	65.8	161	2	Q95UW6	Q95UW6 drosophila
26	196	65.8	255	2	Q9SIH2	Q9SIH2 arabidopsis
27	196	65.8	301	2	O15647	O15647 plasmodium
28	195.5	65.6	192	2	Q92P87	Q92P87 rhizobium m
29	195	65.4	344	2	Q42403	Q42403 gallus gallus
30	194.5	65.3	163	2	Q95UX4	Q95UX4 drosophila
31	194.5	65.3	165	2	Q95NR6	Q95NR6 drosophila

RESULT 4

[illegible]


```

Query Match      67.3%; Score 200.5; DB 2; Length 535;
Best Local Similarity 68.4%; Pred. No. 4.6e-09;
Matches 39; Conservative 0; Mismatches 3; Indels 15; Gaps 2;

Qy 1 GGGGGGGGGGGG-----GGGGGGGGYGRKKRRRRRRGGGGGGGGGGGGGGGGGGGGGG 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 72 GGGGGGGYGGGGGGYGGGGGGGGGGGGY-----GGGGGGRRGGGGGGGGGGGG 119

RESULT 12
Q8QXX8
ID Q8QXX8 PRELIMINARY; PRT; 698 AA.
AC Q8QXX8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE EsV-1-144.
GN Name=ORF 144;
OS Ectocarpus siliculosus virus.
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phaeovirus.
OX NCBI_TaxID=37665;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=EsV-1;
RX MEDLINE=20192171; PubMed=10725207; DOI=10.1006/viro.2000.0225;
RA Delaroue N., Wolf S., Muller D.G., Knippers R.;
RT "Characterization and immunolocalization of major structural proteins
in the brown algal virus EsV-1.";
RL Virology 269:148-155(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=EsV-1;
RA Delaroue N., Bothe G., Pohl T., Knippers R., Mueller D.G., Boland W.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF204951; AAF28325.1; -.
SQ SEQUENCE 698 AA; 74035 MW; C7638DDB922E0D50 CRC64;

Query Match      67.3%; Score 200.5; DB 2; Length 698;
Best Local Similarity 68.4%; Pred. No. 5.8e-09;
Matches 39; Conservative 0; Mismatches 3; Indels 15; Gaps 2;

Qy 1 GGGGGGGGGG-----GGGGGGGGYGRKKRRRRRRGGGGGGGGGGGGGGGGGGGGGG 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 454 GGGGGGGTGGAGGAGGGGGGGGGTGGAG-----GGGGGGGGGGGGGGGGGGGG 501

RESULT 13
Q95UX2
ID Q95UX2 PRELIMINARY; PRT; 165 AA.
AC Q95UX2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE No on or off transient A (fragment).
GN Name=nonA;
OS Drosophila virilis (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=SBB;
RA Huttunen S., Vieira J., Hoikkala A.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY012591; AAG48861.1; -.
FL FlyBase; FBgn0042729; Dvir\nonA.
FT NON TER 1
FT NON TER 165
SQ SEQUENCE 165 AA; 15822 MW; F0D87E1E78DC01F CRC64;

Query Match      66.8%; Score 199; DB 2; Length 165;

```

RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Shydsdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537559;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
 RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.;
 RT "A Drosophila full-length cDNA resource";
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 RN [4]
 RP SEQUENCE OF 5-140 FROM N.A.
 RX MEDLINE=87230988; PubMed=2884623;
 RA Flavell A.J., Dyson J., Ish-Horowicz D.;
 RT "A novel GC-rich dispersed repeat sequence in Drosophila
 RT melanogaster";
 RL Nucleic Acids Res. 15:4035-4048(1987).
 CC -!- FUNCTION: Fibrillar is a component of a nucleolar small nuclear
 CC ribonucleoprotein particle thought to participate in the first
 CC step in processing preribosomal RNA. It is associated with the U3,
 CC U8 and U13 small nuclear RNAs (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear; fibrillar region of the nucleolus
 CC (By similarity).
 CC -!- PTM: By homology to other fibrillins, some or all of the N-
 CC terminal domain arginines are modified to asymmetric
 CC dimethylarginine (DMA) (By similarity).
 CC -!- SIMILARITY: Belongs to the fibrillar family.
 CC -!- CAUTION: Ref.4 sequence differs from that shown due to erroneous
 CC gene model prediction and to a frameshift in position 8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF003459; AAF46950.1; -;
 CC EMBL; E001738; AAN71493.1; -;
 CC EMBL; X05285; CAA28903.1; ALT_FRAME.
 CC HSSP; O57811; 1G8A.
 CC IntAct; Q9W1V3; -;
 CC FlyBase; FBgn003062; Fib.
 CC InterPro; IPR000692; Fibrillar.
 CC Pfam; PF01269; Fibrillar; 1.
 CC PRINTS; PR00052; FIBRILLARIN.
 CC ProDom; PD004637; Fibrillar; 1.
 CC PROSITE; PS00566; FIBRILLARIN; 1.
 CC Methylation; Nuclear protein; Ribonucleoprotein; RNA-binding;
 KW rRNA processing.
 FT DOMAIN 9 108 Gly/DNA-rich.
 FT MOD_RES 9 9 Asymmetric dimethylarginine (By
 FT similarity).
 FT MOD_RES 23 23 Asymmetric dimethylarginine (By
 FT similarity).
 FT MOD_RES 25 25 Asymmetric dimethylarginine (By
 FT similarity).
 FT MOD_RES 40 40 Asymmetric dimethylarginine (By
 FT similarity).
 FT MOD_RES 42 42 Asymmetric dimethylarginine (By
 FT similarity).
 FT MOD_RES 48 48 Asymmetric dimethylarginine (By
 FT similarity).
 FT MOD_RES 51 51 Asymmetric dimethylarginine (Probable).
 FT MOD_RES 58 58 Asymmetric dimethylarginine (Probable).

FT MOD_RES 63 63 Asymmetric dimethylarginine (Probable).
 FT MOD_RES 71 71 Asymmetric dimethylarginine (Probable).
 FT MOD_RES 77 77 Asymmetric dimethylarginine (Probable).
 FT MOD_RES 83 83 Asymmetric dimethylarginine (Probable).
 FT MOD_RES 88 88 Asymmetric dimethylarginine (Probable).
 FT MOD_RES 93 93 Asymmetric dimethylarginine (Probable).
 FT MOD_RES 98 98 Asymmetric dimethylarginine (Probable).
 FT CONFLICT 55 67 Missing (in Ref. 3).
 FT CONFLICT 134 134 R -> G (in Ref. 3).
 SQ SEQUENCE 344 AA; 34637 MW; 58B536FAACAE01D6 CRC64;
 Query Match 66.8%; Score 199; DB 1; Length 344;
 Best Local Similarity 65.1%; Pred. No. 4.2e-09;
 Matches 41; Conservative 1; Mismatches 5; Indels 16; Gaps 3;
 QY 1 GGGGGGGGGG---GGGGGGGGGGVGRKRRRRRRGGG-----GGGGGGGGGGGG 48
 DB 12 GGGGGGGGGGFRGCGGGGGGGGGGFG----GGRGRRGGGDRGGRGGGGRGGGG 67
 QY 49 GGG 51
 DB 68 GGG 70
 RESULT 15
 FBRL_DROER STANDARD; PRT; 345 AA.
 AC Q811F4;
 DT 25-JAN-2005 (Rel. 46, Created)
 DT 25-JAN-2005 (Rel. 46, Last sequence update)
 DT 25-JAN-2005 (Rel. 46, Last annotation update)
 DE Fibrillar.
 GN Name=Fib;
 OS Drosophila erecta (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426072; PubMed=12537575;
 RA Bergman C.M., Pfeiffer B.D., Rincon-Limas D.E., Hoskins R.A.,
 RA Gnirke A., Mungall C.J., Wang A.M., Krommiller B., Pacleb J.M.,
 RA Park S., Stapleton M., Wan K.H., George R.A., de Jong P.J., Botas J.,
 RA Rubin G.M., Celniker S.E.;
 RT "Assessing the impact of comparative genomic sequence data on the
 RT functional annotation of the Drosophila genome";
 RL Genome Biol. 3:RESEARCH0086.1-RESEARCH0086.20(2002).
 CC -!- FUNCTION: Fibrillar is a component of a nucleolar small nuclear
 CC ribonucleoprotein particle thought to participate in the first
 CC step in processing preribosomal RNA. It is associated with the U3,
 CC U8 and U13 small nuclear RNAs (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear; fibrillar region of the nucleolus
 CC (By similarity).
 CC -!- PTM: By homology to other fibrillins, some or all of the N-
 CC terminal domain arginines are modified to asymmetric
 CC dimethylarginine (DMA) (By similarity).
 CC -!- SIMILARITY: Belongs to the fibrillar family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AY190941; AAO1021.1; -;
 CC HSSP; O57811; 1G8A.
 CC FlyBase; FBgn0064623; Deref.Fib.
 CC InterPro; IPR000692; Fibrillar.
 CC Pfam; PF01269; Fibrillar; 1.
 CC PRINTS; PR00052; FIBRILLARIN.

```
DR ProDom: PD004637; Fibrillarlin; 1.
KW PROSITE; PS00566; FIBRILLARIN; 1.
KW Methylation; Nuclear protein; Ribonucleoprotein; RNA-binding;
KW rRNA processing.
FT DOMAIN 9 109 Gly/DMA-rich.
FT MOD_RES 9 9 Asymmetric dimethylarginine (By
FT FT similarity).
FT MOD_RES 23 23 Asymmetric dimethylarginine (By
FT FT similarity).
FT MOD_RES 25 25 Asymmetric dimethylarginine (By
FT FT similarity).
FT MOD_RES 41 41 Asymmetric dimethylarginine (By
FT FT similarity).
FT MOD_RES 43 43 Asymmetric dimethylarginine (By
FT FT similarity).
FT MOD_RES 49 49 Asymmetric dimethylarginine (By
FT FT similarity).
FT MOD_RES 52 52 Asymmetric dimethylarginine (Probable).
FT MOD_RES 59 59 Asymmetric dimethylarginine (Probable).
FT MOD_RES 64 64 Asymmetric dimethylarginine (Probable).
FT MOD_RES 72 72 Asymmetric dimethylarginine (Probable).
FT MOD_RES 78 78 Asymmetric dimethylarginine (Probable).
FT MOD_RES 84 84 Asymmetric dimethylarginine (Probable).
FT MOD_RES 89 89 Asymmetric dimethylarginine (Probable).
FT MOD_RES 94 94 Asymmetric dimethylarginine (Probable).
FT MOD_RES 99 99 Asymmetric dimethylarginine (Probable).
SQ SEQUENCE 345 AA; 34709 MW; F05DAB01491AE66f CRC64;

Query Match 66.8%; Score 199; DB 1; Length 345;
Best Local Similarity 65.1%; Pred. No. 4.3e-09;
Matches 41; Conservative 1; Mismatches 5; Indels 16; Gaps 3;

QY 1 GGGGGGGGGG----GGGGGGGGGGGGYGRKKRRQRRRGGG-----GGGGGGGGGGGGG 48
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 13 GGGGGGGGGGFRGRGGGGGGGGGGGFG----GGRGRGGGGDRGGRGFRGGGGGGGGG 68
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 49 GGG 51
    |||
Db 69 GGG 71
```

Search completed: September 7, 2005, 12:59:58
Job time : 113.5 secs

This Page Blank (uspio)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	298	100.0	51	7	ADC65108	Adc65108 HIV-TAT p
2	277	93.0	51	7	ADC65109	Adc65109 HIV-TAT p
3	230.5	77.3	920	7	ADC87061	Adc87061 Human GPC
4	229.5	77.0	200	8	ADG88856	Adg88856 Poly Gly
5	229.5	77.0	200	8	ADJ84541	Adj84541 T2R G-pro
6	229.5	77.0	200	8	ADM96215	Adm96215 Poly Gly
7	229.5	77.0	201	5	ABG71303	Abg71303 PINPOINT
8	210.5	70.6	56	5	ABF29010	Abf29010 Streptococ
9	206.5	69.3	399	4	ABR60010	AbR60010 Streptococ
10	203	68.1	188	6	ABU22369	Abu22369 Protein e
11	198.5	66.6	208	8	ABO59421	AbO59421 Human gen
12	197.5	66.3	237	6	ABU40120	Abu40120 Protein e
13	197	66.1	344	4	ABB64729	Abb64729 Drosophil
14	193	64.8	258	3	AA551723	Aa551723 Arabidops
15	192.5	64.6	38	7	ADH22061	Adh22061 Synthetic
16	191	64.1	345	8	ADK14010	Adk14010 Chicken m
17	189.5	63.6	462	8	ADE77159	AdE77159 Human pro
18	189	63.4	283	4	AAm17023	Aam17023 Peptide #
19	189	63.4	283	4	ABB36017	Abb36017 Peptide #
20	189	63.4	283	4	AAm29514	Aam29514 Peptide #
21	189	63.4	283	4	ABB30839	Abb30839 Peptide #
22	189	63.4	283	4	ABB21422	Abb21422 Protein #
23	189	63.4	283	4	AAm69190	Aam69190 Human bon
24	189	63.4	283	4	AAm56810	Aam56810 Human bra
25	189	63.4	283	4	ABG50867	AbG50867 Human liv

CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins (AB57737-
 CC AB572072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 399 AA;
 Query Match 69.3%; Score 206.5; DB 4; Length 399;
 Best Local Similarity 68.9%; Pred. No. 6.7e-12;
 Matches 42; Conservative 0; Mismatches 2; Indels 17; Gaps 3;
 QY 1 GGGGGGGG-----GGGGGGGGGGGRRDRRRGGGGGGG-----GGGGGGGGGG 48
 Db 215 GGGGGGGGGGFGGRRGGGGGGGGGGGGGRRF-----RGGGGGGGRRYDRGGGGGGGGG 269
 QY 49 G 49
 Db 270 G 270
 RESULT 10
 ABU22369
 ID ABU22369 standard; protein; 188 AA.
 XX
 AC ABU22369,
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #7896.
 DE
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Burkholderia mallei.
 OS
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 23-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362659P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlssen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA26239.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 50293; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 188 AA;
 Query Match 68.1%; Score 203; DB 6; Length 188;
 Best Local Similarity 68.9%; Pred. No. 7.9e-12;
 Matches 42; Conservative 0; Mismatches 9; Indels 10; Gaps 3;
 QY 1 GGGGGGGGGGGG--GGGGGGGGRRDRRRGGGGG---GGGG-----GGGGGGGGGG 50
 Db 116 GGGGGGGGGGGGDDGGGGGGGGYGGGRDMMRGGGGGGRASGGGGGGGGGGG 175
 QY 51 G 51
 Db 176 G 176
 RESULT 11
 ABO59421
 ID ABO59421 standard; protein; 208 AA.
 XX
 AC ABO59421;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon protein #5655.
 XX
 KW Human; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 OS Homo sapiens.
 XX
 PN US2003194704-A1.
 XX
 PD 16-OCT-2003.
 XX
 XX 03-APR-2002; 2002US-00029386.
 PF
 XX 03-APR-2002; 2002US-00029386.
 PR
 PA (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX
 XX Penn SG, Rank DR, Hanzel DK;
 PI WPI; 2004-119264/12.
 DR
 XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.

OS Drosophila melanogaster.
XX WO200171042-A2.
XX
XX
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL08832.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Disclosure; SEQ ID NO 20979; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABU16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 344 AA;
SQ

Query Match 66.1%; Score 197; DB 4; Length 344;
Best Local Similarity 70.9%; Pred. No. 4.6e-11;
Matches 39; Conservative 0; Mismatches 12; Indels 4; Gaps 1;

Qy 1 GGGGGG---GGGGGGGGGGGGRRRRRRRRGGGGGGGGGGGGGGGGGG 51
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 16 GGGGGGPRGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 70
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 14
AAG51723
ID AAG51723 standard; protein; 258 AA.
XX
XX AAG51723;
AC
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 65674.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX
XX 05-MAR-1999; 99US-0123180P.
XX
XX 09-MAR-1999; 99US-0123548P.
XX
XX 23-MAR-1999; 99US-0125788P.
XX
XX 23-MAR-1999; 99US-0126264P.
XX

PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.


```

PR 27-MAR-2002; 2002US-0367500P.
PR 08-APR-2002; 2002US-0370227P.
PR 10-MAY-2002; 2002US-0378950P.
PR 24-JUL-2002; 2002US-0398008P.
PR 09-AUG-2002; 2002US-0402131P.
PR 13-AUG-2002; 2002US-0402708P.
PR 18-SEP-2002; 2002US-0411355P.
PR 02-OCT-2002; 2002US-0414984P.
PR 11-OCT-2002; 2002US-0417611P.
PR 23-OCT-2002; 2002US-0420246P.
PR 05-NOV-2002; 2002US-0423623P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PR PI Rosen CA, Haseltine WA;
XX
XX WPI; 2003-598501/56.
XX
PT New albumin fusion protein, useful for preparing a composition for
XX treating diabetes mellitus.
XX
XX Example 24; SEQ ID NO 858; 1086pp; English.
XX
XX The invention relates to fusion proteins comprising human serum albumin
XX (ADH21530) and a therapeutic polypeptide such as a therapeutic protein,
XX antibody or peptide or their variants or fragments. The therapeutic
XX protein may be fused to the N-terminus, the C-terminus or both termini of
XX albumin via a linker. The albumin component of the fusion proteins
XX prolongs the shelf-life and the in vitro and vivo biological activity of
XX the proteins compared with those of the corresponding therapeutic
XX proteins on their own. The invention also relates to nucleic acids
XX encoding albumin fusion proteins, vectors and host cells comprising an
XX albumin fusion protein nucleic acid, compositions and kits comprising an
XX albumin fusion protein, the method of extending the shelf-life of a
XX therapeutic protein by fusion with albumin, and the treatment of disease
XX using an albumin fusion protein. The albumin fusion proteins may be used
XX in the treatment of metabolic/endocrine disorders, diabetes and diabetes-
XX related conditions. Specifically the albumin fusion proteins may be used
XX to treat type 1 and type 2 diabetes, hyperglycaemia, neural disorders
XX (especially neuropathy), retinopathy, cardiovascular disorders
XX (especially heart disease, renal disorders and obesity. The proteins may
XX also be used in a method of maintaining a basal glucose level in a
XX patient and in a method for losing weight. The present sequence is
XX related to the invention.
XX
SQ Sequence 38 AA;
XX
XX Query Match 64.6%; Score 192.5; DB 7; Length 38;
XX Best Local Similarity 74.5%; Pred. No. 2.1e-11;
XX Matches 35; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
XX
QY 1 GGGGGGGGGGGGGGGGGGGGRDRRRRRGGGGGGGGGGGGGGGG 47
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 GGGGGGGGGGGGGGGGGGGG-----GGGGGGGGGGGGGGPG 36
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX
Search completed: September 7, 2005, 12:56:07
Job time : 68.5 secs

```



```

; REFERENCE/DOCKET NUMBER: 018781-0008000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1..97
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "gly at positions 1-97 may be
; OTHER INFORMATION: present or absent"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 105..201
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "gly at positions 105-201 may be
; OTHER INFORMATION: present or absent"
; US-09-053-003-40
; Query Match 77.0%; Score 229.5; DB 3; Length 201;
; Best Local Similarity 80.4%; Pred.No.3.8e-15;
; Matches 41; Conservative 0; Mismatches 1; Indels 9;
; QY 1 GGGGGGGGGGGGGGGGGGRDRRRRQRRGGGGGGGGGGGGGGGGGGGGG 51
; Db 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 91

```

RESULT 4
 US-09-054-281-22
 ; Sequence 22, Application US/09054281
 ; Patent No. 644421
 ; GENERAL INFORMATION:
 ; APPLICANT: Chung, Jay H.
 ; TITLE OF INVENTION: Methods for Detecting Intermolecular
 ; TITLE OF INVENTION: Interactions In Vivo and In Vitro
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/054,281
 ; FILING DATE: 02-APR-1998
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/080,234
 ; FILING DATE: 03-APR-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/825,622
 ; FILING DATE: 03-APR-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith, Timothy L.
 ; REGISTRATION NUMBER: 35,367
 ; REFERENCE/DOCKET NUMBER: 015280-295200US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 22:

[illegible]

RESULT 11

US-09-575-574-4
; Sequence 4, Application US/09575574
; Patent No. 6376750
; GENERAL INFORMATION:
; APPLICANT: YU, Su-May
; APPLICANT: CHAO, Yu-Chan
; TITLE OF INVENTION: PLANT SEEDLING AND EMBRYO PROMOTER
; FILE REFERENCE: 08919-047001
; CURRENT APPLICATION NUMBER: US/09/575,574
; CURRENT FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-575-574-4

Query Match 62.4%; Score 186; DB 3; Length 162;
Best Local Similarity 72.0%; Pred. No. 4e-11;
Matches 36; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GGGGGGGGGGGGGGGGGGGRRDRRRRRGGGGGGGGGGGGGGGGGGGGGG 50
Db 97 GGGGGGGGGGGGGGGGGGGYGGGGGGYGGRRGGYGGGGGGGGGGGGGG 146

RESULT 12

US-09-832-297A-12
; Sequence 12, Application US/09832297A
; Patent No. 6652836
; GENERAL INFORMATION:
; APPLICANT: LUIGEN, George A.
; APPLICANT: FluoroProbe, Inc.
; TITLE OF INVENTION: METHOD FOR VIEWING TUMOR TISSUE LOCATED WITHIN A BODY CAVITY
; FILE REFERENCE: FLUOR1120-2
; CURRENT APPLICATION NUMBER: US/09/832,297A
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/362,805
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/173,190
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Peptide linker moiety
; NAME/KEY: REPEAT
; LOCATION: (3)..(6)
; OTHER INFORMATION: Amino Acid at residue 3 could be repeated up to 3 times
; NAME/KEY: REPEAT
; LOCATION: (3)..(57)
; OTHER INFORMATION: Amino Acids at residues 3 to 7 could be repeated up to 10 times

US-09-832-297A-12
Query Match 61.4%; Score 183; DB 4; Length 60;
Best Local Similarity 67.3%; Pred. No. 3.4e-11;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
Qy 2 GGGGGGGGGGGGGGGGGGGRRDRRRRRGGGGGGGGGGGGGGGGGGGGGG 50
Db 3 GGG 51

RESULT 13

US-09-411-067C-4

; Sequence 4, Application US/09411067C

; Patent No. 6578610
; GENERAL INFORMATION:
; APPLICANT: NUVAS, LLC
; APPLICANT: HOUSTON, L.L.
; TITLE OF INVENTION: USE OF A CONTEXT-DEPENDENT FUNCTIONAL ENTITY TO ENHANCE THE EFFICACY OF AN AGENT
; FILE REFERENCE: NOVAS1140
; CURRENT APPLICATION NUMBER: US/09/411,067C
; CURRENT FILING DATE: 2002-05-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Spacer element
; NAME/KEY: REPEAT
; LOCATION: (1)..(5)
; OTHER INFORMATION: Spacer peptide variation; sequence repeated "n" times, where n = 1-20
; OTHER INFORMATION: 1-20
US-09-411-067C-4

Query Match 61.4%; Score 183; DB 4; Length 100;
Best Local Similarity 67.3%; Pred. No. 5.2e-11;
Matches 33; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 2 GGGGGGGGGGGGGGGGGGGRRDRRRRRGGGGGGGGGGGGGGGGGGGGGG 50
Db 1 GGG 49

RESULT 14

US-08-864-038A-3
; Sequence 3, Application US/08864038A
; Patent No. 6001592

; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY TO SAID POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: 812-5 Hirano
; STREET: Isshinden
; CITY: Tsu-city
; STATE: Mie-prefecture
; COUNTRY: JAPAN
; ZIP: 514-01
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Microsoft Windows 95
; SOFTWARE: Word Perfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/864,038A
; FILING DATE: May 28, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-184459
; FILING DATE: 15-July-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5610
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

GenCore version 5.1.6
Copyright (c) 1995 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2005, 12:54:00 ; Search time 57 Seconds
(without alignments)
352.368 Million cell updates/sec

Title: US-09-910-432-19

Perfect score: 298

Sequence: 1 GGGGGGGGGGGGGGGGGG.....GGGGGGGGGGGGGGGGGG 51

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep.*

2: /cgn2_6/ptodata/1/pubaa/PCT_NEW_PUB.pep.*

3: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*

4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep.*

5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep.*

6: /cgn2_6/ptodata/1/pubaa/PCTUS_PUBCOMB.pep.*

7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep.*

8: /cgn2_6/ptodata/1/pubaa/US08_PUBCOMB.pep.*

9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep.*

10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep.*

11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep.*

12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUB.pep.*

13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep.*

14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep.*

15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep.*

16: /cgn2_6/ptodata/1/pubaa/US10D_PUBCOMB.pep.*

17: /cgn2_6/ptodata/1/pubaa/US10E_PUBCOMB.pep.*

18: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep.*

19: /cgn2_6/ptodata/1/pubaa/US11A_PUBCOMB.pep.*

20: /cgn2_6/ptodata/1/pubaa/US11_NEW_PUB.pep.*

21: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep.*

22: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	298	100.0	51	10	US-09-910-432-19
2	277	93.0	51	10	US-09-910-432-20
3	230.5	77.3	920	15	US-10-292-798-1514
4	229.5	77.0	200	9	US-09-798-584-18
5	229.5	77.0	200	9	US-09-967-624-19
6	229.5	77.0	200	9	US-09-998-667-18
7	229.5	77.0	200	10	US-09-921-159-34
8	229.5	77.0	200	10	US-09-990-940-21
9	229.5	77.0	200	10	US-09-989-981A-13
10	229.5	77.0	200	10	US-09-850-948-29
11	229.5	77.0	200	14	US-10-160-354-4
12	229.5	77.0	200	14	US-10-026-331B-21
13	229.5	77.0	200	14	US-10-026-021-8
14	229.5	77.0	200	14	US-10-161-165-3
15	229.5	77.0	200	14	US-10-160-663-3
16	229.5	77.0	200	14	US-10-071-838-15
17	229.5	77.0	200	14	US-10-094-417-25
18	229.5	77.0	200	14	US-10-188-405-13
19	229.5	77.0	200	14	US-10-273-575-29
20	229.5	77.0	200	14	US-10-233-098-5
21	229.5	77.0	200	14	US-10-245-850-3
22	229.5	77.0	200	14	US-10-293-582-27
23	229.5	77.0	200	14	US-10-100-818-14
24	229.5	77.0	200	14	US-10-237-467-18
25	229.5	77.0	200	14	US-10-321-204-54
26	229.5	77.0	200	14	US-10-179-766-12
27	229.5	77.0	200	14	US-10-123-568-4
28	229.5	77.0	200	14	US-10-123-731-10
29	229.5	77.0	200	15	US-10-308-393-53
30	229.5	77.0	200	15	US-10-352-724-5
31	229.5	77.0	200	15	US-10-328-916-57
32	229.5	77.0	200	15	US-10-452-015-3
33	229.5	77.0	200	15	US-10-364-861-94
34	229.5	77.0	200	15	US-10-231-956A-521
35	229.5	77.0	200	15	US-10-255-775-3
36	229.5	77.0	200	15	US-10-459-190-24
37	229.5	77.0	200	15	US-10-339-744-4
38	229.5	77.0	200	16	US-10-649-400-7
39	229.5	77.0	200	16	US-10-620-052A-78
40	229.5	77.0	200	16	US-10-794-897-9
41	229.5	77.0	200	17	US-10-616-403-7
42	229.5	77.0	200	17	US-10-679-102-32
43	229.5	77.0	200	17	US-10-696-909A-72
44	229.5	77.0	200	18	US-10-491-545A-96
45	229.5	77.0	200	18	US-10-805-075-10

ALIGNMENTS

RESULT 1

US-09-910-432-19

; Sequence 19, Application US/09910432

; Publication No. US20030229034A1

; GENERAL INFORMATION:

; APPLICANT: Waugh, Jacob

; APPLICANT: Dake, Michael

; APPLICANT: Essentia Biosystems, Inc.

; TITLE OF INVENTION: Multi-Component Biological Transport Systems

; FILE REFERENCE: 020154-000110US

; CURRENT APPLICATION NUMBER: US/09/910,432

; CURRENT FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 60/220,244

; PRIOR FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 19

; TYPE: PRT

; LENGTH: 51

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:positively charged HIV-TAR fragment branching group

; OTHER INFORMATION: (efficiency group) attached to solid backbone

; FEATURE:

; NAME/KEY: MOD RES

; LOCATION: (1)..(20)

; OTHER INFORMATION: Gly at positions 1-20 may be present or absent

; FEATURE:

; NAME/KEY: MOD RES

; LOCATION: (32)..(51)

; OTHER INFORMATION: Gly at positions 32-51 may be present or absent

US-09-910-432-19

Sequence 21, Appli

Sequence 8, Appli

Sequence 3, Appli

Sequence 3, Appli

Sequence 15, Appli

Sequence 25, Appli

Sequence 13, Appli

Sequence 29, Appli

Sequence 5, Appli

Sequence 27, Appli

Sequence 14, Appli

Sequence 18, Appli

Sequence 54, Appli

Sequence 12, Appli

Sequence 10, Appli

Sequence 53, Appli

Sequence 5, Appli

Sequence 57, Appli

Sequence 94, Appli

Sequence 521, Appli

Sequence 3, Appli

Sequence 24, Appli

Sequence 4, Appli

Sequence 7, Appli

Sequence 9, Appli

Sequence 7, Appli

Sequence 32, Appli

Sequence 72, Appli

Sequence 96, Appli

Sequence 10, Appli


```

RESULT 6
US-09-998-667-18
; Sequence 18, Application US/09998667
; Patent No. US20020146747A1
; GENERAL INFORMATION:
; APPLICANT: Masuda, Esteban
; APPLICANT: Liao, X. Charlene
; APPLICANT: Zhao, Haozan
; APPLICANT: Chu, Peter
; APPLICANT: Pardo, Jorge
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: FRACs: Modulators of Lymphocyte Activation
; FILE REFERENCE: 021044-0006000US
; CURRENT APPLICATION NUMBER: US/09/998,667
; CURRENT FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: US 60/282,432
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 18
; LENGTH: 200
; TYPE: PRT

```

```

RESULT 8
US-09-990-940-21
; Sequence 21, Application US/09909040
; Publication No. US2003002752A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: An, Songzhu
; APPLICANT: Dai, Kang
; APPLICANT: Gupte, Jamila S.
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US2003002752A1el Receptors
; FILE REFERENCE: 018781-0074100S
; CURRENT APPLICATION NUMBER: US/09/990,940
; CURRENT FILING DATE: 2001-11-21

```



```

; OTHER INFORMATION: Description of Artificial Sequence:poly Gly tag
; OTHER INFORMATION: flexible linker
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (6)..(200)
; OTHER INFORMATION: Gly residues from position 6 to 200 may be present
; OTHER INFORMATION: or absent
US-10-160-663-3

Query Match      77.0%; Score 229.5; DB 14; Length 200;
Best Local Similarity 80.4%; Pred. No. 6.7e-12;
Matches 41; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

Qy      1 GGGGGGGGGGGGGGGGGGGGGRRRRGGGGGGGGGGGGGGGGGGGG 51
      |||||
Db      1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 42
      |||||

```

Search completed: September 7, 2005, 13:03:57
Job time : 58 secs

This Page Blank (uspio)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	223	74.8	221	2	T04592	glycine-rich cell	
2	212.5	71.3	404	2	S54729	RNA-binding protein	
3	206	69.1	136	2	T29282	hypothetical protein	
4	206	69.1	185	1	KNR251	glycine-rich cell	
5	203	68.1	285	2	T31503	hypothetical protein	
6	201	67.4	407	2	T07381	glycine-rich protein	
7	198.5	66.6	440	2	S71795	transcription factor	
8	195	65.4	1236	2	T24045	hypothetical protein	
9	194.5	65.3	106	2	F84797	hypothetical protein	
10	194.5	65.3	185	2	T03583	glycine-rich RNA-b	
11	194	65.1	396	2	T49109	glycine-rich protein	
12	192.5	64.6	385	2	T20410	hypothetical protein	
13	192	64.4	255	2	B84777	hypothetical protein	
14	191	64.1	239	2	S49193	GCR 101 protein -	
15	190.5	63.9	199	2	S16063	acp-22 protein - Y	
16	190.5	63.9	199	2	S32224	acp-22 protein - Y	
17	190	63.8	291	1	S31415	glycine-rich protein	
18	187.5	62.9	892	2	T27005	hypothetical protein	
19	186.5	62.6	384	1	A26099	glycine-rich cell	
20	186.5	62.6	462	4	S33798	FUS/CHOP mutant fu	
21	186.5	62.6	526	1	S33799	RNA-binding protein	
22	186	62.4	142	2	S12311	glycine-rich RNA-b	
23	186	62.4	528	2	G02127	fus-like protein -	
24	185.5	62.2	168	1	S12312	glycine-rich RNA-b	
25	185.5	62.2	302	2	K84470	hypothetical protein	
26	185	62.1	214	1	KNNT28	glycine-rich protein	
27	184.5	61.9	393	2	T20268	hypothetical protein	
28	184	61.7	171	2	H84709	probable glycine-r	
29	183.5	61.6	167	2	S71779	glycine-rich RNA-b	

A;Residues: 1-106 <STO>
A;Cross-references: UNIPROT:Q7XJP7; GB:AE002093; NID:g4895201; PIDN:AAD32788.1; GSPDB:GN0062;
C;Genetics:
A;Gene: AT2g37830
A;Map position: 2
C;Superfamily: Arabidopsis glycine-rich protein 3

Query Match 65.3%; Score 194.5; DB 2; Length 106;
Best Local Similarity 55.7%; Pred. No. 1.5e-09;
Matches 39; Conservative 0; Mismatches 12; Indels 19; Gaps 1;

QY 1 GGGGGGGGGGGGGGGGGGGRRD-----DRRQRRRRGGGGGGGGG 41
DB 31 GGG 90
QY 42 GGGGGGGGGGG 51
DB 91 GGGGGGGGGGG 100

RESULT 10
T03583
glycine-rich RNA-binding protein - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03583
R;Lee, M.C.; Kim, C.S.; Yi, B.Y.; Lee, J.S.; Eun, M.Y.
submitted to the EMBL Data Library, June 1997
A;Description: Isolation and characterization of RNA-binding glycine rich protein of rice.
A;Reference number: Z14958
A;Accession: T03583
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-165 <LEE>
A;Cross-references: UNIPROT:Q24184; EMBL:AF009411; NID:g2267592; PIDN:AAB63589.1; PID:g2267592;
C;Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology
F;9-76/Domain: ribonucleoprotein repeat homology <RRM>

Query Match 65.3%; Score 194.5; DB 2; Length 165;
Best Local Similarity 67.2%; Pred. No. 2.1e-09;
Matches 41; Conservative 0; Mismatches 5; Indels 15; Gaps 3;

QY 1 GGGGGGGG---GGGGG-----GGGGGGGGRRDRRQRRRRGGGGGGGGGGGGGGGGGG 50
DB 90 GGGGGGGGQRRGG 144
QY 51 G 51
DB 145 G 145

RESULT 11
T49109
glycine-rich protein - Arabidopsis thaliana
N;Alternate names: protein AT4g22020
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C;Accession: T49109
R;Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.; M.
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25016
A;Accession: T49109
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-396 <BEV>
A;Cross-references: UNIPROT:O65450; EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g22020
A;Experimental source: cultivar Columbia; BAC clone F1N20
C;Genetics:
A;Gene: ATSP:AT4g22020
A;Map position: 4
C;Superfamily: Phaseolus glycine-rich cell wall protein 1.8

